

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-301880
 (43)Date of publication of application : 13.11.1998

(51)Int.Cl. G06F 13/00
 G06F 12/00
 G06F 15/163

(21)Application number : 10-091588
 (22)Date of filing : 03.04.1998

(71)Applicant : AT & T CORP
 (72)Inventor : CHIANG JOHN J
 DUTTA PARTHA P
 LONDON THOMAS B
 SRBLJIC SINISA
 VRSALOVIC DALIBOR F

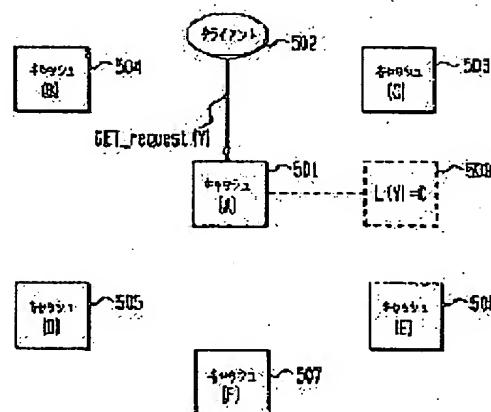
(30)Priority

Priority number : 97 827763 Priority date : 10.04.1997 Priority country : US

(54) SCALEABLE NETWORK OBJECT CACHE

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a distributed cache system for a network with excellent scalability.
SOLUTION: Object caches 503-507 are arranged on a network, and the copy of a data object from an original information source is stored in the several object caches. When a request to a certain object is issued from a client 502 to a receiving cache 501, which object cache 503-507 stores a directory list related with the object is searched by executing a directory locator function 508, and the receiving cache 501 inquires where the copy of the data object is stored in the searched cache.



LEGAL STATUS

[Date of request for examination] 12.02.2002

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

(19) 日本国特許庁(JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-301880

(43) 公開日 平成10年(1998)11月13日

(51) Int. C1. 6

G 06 F 13/00 3 5 5
12/00 5 4 6
15/163

識別記号

F I

G 06 F 13/00 3 5 5
12/00 5 4 6 K
15/16 3 2 0 K

審査請求 未請求 請求項の数 20 O L

(全 22 頁)

(21) 出願番号 特願平10-91588

(22) 出願日 平成10年(1998)4月3日

(31) 優先権主張番号 08/827,763

(32) 優先日 1997年4月10日

(33) 優先権主張国 米国 (U S)

(71) 出願人 390035493

エイ・ティ・アンド・ティ・コーポレーション
AT&T CORP.

アメリカ合衆国 10013-2412 ニューヨーク ニューヨーク アヴェニュー オブ ジ アメリカズ 32

(72) 発明者 ジョン ジェイ チャン

アメリカ合衆国 カリフォルニア州 サン フランシスコ 21エスティ アヴェニュー 1630

(74) 代理人 弁理士 吉田 研二 (外2名)

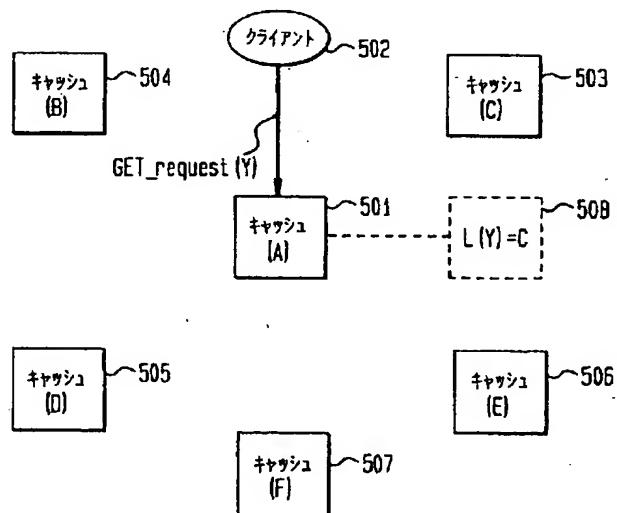
最終頁に続く

(54) 【発明の名称】スケーラブルネットワークオブジェクトキャッシュ

(57) 【要約】

【課題】 スケーラビリティに優れたネットワーク用分散型キャッシュシステムを提供する。

【解決手段】 ネットワーク上にオブジェクトキャッシュ 503～507 を配置しておき、そのいくつかにオリジナル情報源からのデータオブジェクトのコピーを記憶しておく。受信キャッシュ 501 にクライアント 502 からあるオブジェクトに対するリクエストがあれば、そのオブジェクトに関するディレクトリリストがどのオブジェクトキャッシュ 503～507 に記憶されているかをディレクトリロケータ関数 508 を実行することにより求め、受信キャッシュ 501 はその求められたキャッシュにデータオブジェクトのコピーが何処に記憶されているかを問い合わせる。



【特許請求の範囲】

【請求項1】 ネットワークにおける分散型キャッシュシステムであって、

プロセッサとコンピュータにより読み出し可能なメモリとネットワークに接続するポートとを接続するデータバスを含み、ユーザからオブジェクトに対するリクエストを受信し、前記リクエストを入力として用いてポイントを出力として与えるディレクトリロケータ関数を実行する、受信キャッシュと、

プロセッサとコンピュータにより読み出し可能なメモリとネットワークに接続するポートとを接続するデータバスを含み、前記ポイントがネットワークアドレスを示す場合に前記ポイントによってそのネットワークアドレスが示され、前記メモリ中にリクエストされた前記オブジェクトのディレクトリリストを記憶する、ディレクトリキャッシュと、を含み、

前記ディレクトリリストは、リクエストされた前記オブジェクトのネットワークアドレスと、ネットワーク上のオブジェクトキャッシュのネットワークアドレスとを含み、

前記オブジェクトキャッシュは、プロセッサとコンピュータにより読み出し可能なメモリとネットワークに接続するポートとを接続するデータバスを含み、前記メモリ中にリクエストされた前記オブジェクトのコピーを記憶するシステム。

【請求項2】 請求項1に記載のシステムにおいて、前記受信キャッシュは、前記ポイントがネットワークアドレスを示さない場合、前記ユーザによってリクエストされたオブジェクトに対するリクエストを前記オブジェクトのネットワークアドレスに送信するシステム。

【請求項3】 請求項1に記載のシステムにおいて、前記受信キャッシュ、前記ディレクトリキャッシュ、および前記オブジェクトキャッシュは、ネットワーク上の階層レベルの第1レベルで論理上グループ分けされ、前記システムはさらに、前記階層レベルの第2レベルで論理上グループ分けされる第2の受信キャッシュと、第2のディレクトリキャッシュと、第2のオブジェクトキャッシュとを含み、

前記第1レベルの前記受信キャッシュは、前記ポイントがネットワークアドレスを示さない場合、前記ユーザからリクエストされたオブジェクトに対するリクエストを前記第2レベルの前記第2の受信キャッシュに送信するシステム。

【請求項4】 ネットワーク上に記憶されたデータオブジェクトに対するユーザからのリクエストを受信する、ネットワーク上のスケーラブルな分散型キャッシングシステムを実現する方法であって、

データオブジェクトのコピーをネットワーク上のオブジェクトキャッシュに記憶するステップと、

ディレクトリリストをネットワーク上のディレクトリキ

ヤッシュに記憶するステップであって、前記ディレクトリリストは、前記オブジェクトのネットワークアドレスと、前記オブジェクトのコピーを記憶していると称されるオブジェクトキャッシュのネットワークアドレスとを含むステップと、

ユーザからオブジェクトに対するリクエストを受信する受信キャッシュ上でロケータ関数を実行するステップであって、前記リクエストは前記オブジェクトのネットワークアドレスを含み、前記ロケータ関数は前記ユーザリクエストを入力として使用してポイントを出力として与えるステップと、を含む方法。

【請求項5】 請求項4に記載の方法において、前記ポイントはネットワークアドレスを示さない方法。

【請求項6】 請求項5に記載の方法であって、前記受信キャッシュから前記オブジェクトの前記ネットワークアドレスにおけるサーバに、前記オブジェクトのコピーを要求するメッセージを送信するステップと、前記オブジェクトの前記ネットワークアドレスにおける前記サーバから前記受信キャッシュに、前記オブジェクトのコピーを送信するステップと、をさらに含む方法。

【請求項7】 請求項6に記載の方法であって、前記受信キャッシュに前記オブジェクトのコピーを記憶するステップと、

前記受信キャッシュから前記ディレクトリキャッシュに、前記オブジェクトのディレクトリキャッシュ上の前記オブジェクトキャッシュのアドレスを削除し、前記受信キャッシュのアドレスを前記ディレクトリリストに加えることを要求するメッセージを送信するステップと、前記オブジェクトの前記ディレクトリキャッシュ上の前記オブジェクトキャッシュのアドレスを削除し、前記受信キャッシュのアドレスを前記ディレクトリリストに加えるステップと、をさらに含む方法。

【請求項8】 請求項5に記載の方法において、前記受信キャッシュは、ネットワーク上のキャッシュの論理レベルの階層中の第1の論理レベルにあり、前記方法は、前記第1レベルの前記受信キャッシュから第2の階層レベルの受信キャッシュに、前記オブジェクトのコピーを要求するメッセージを送信するステップをさらに含む方法。

【請求項9】 請求項4に記載の方法において、前記ポイントは、前記オブジェクトのディレクトリリストを記憶しているディレクトリキャッシュのネットワークアドレスを指す方法。

【請求項10】 請求項9に記載の方法であって、前記受信キャッシュから前記ポイントによって示されたアドレスをもつ前記ディレクトリキャッシュに、オブジェクトリクエストメッセージを送信するステップをさらに含む方法。

【請求項11】 請求項10に記載の方法であって、前記ディレクトリキャッシュから前記受信キャッシュに、

前記オブジェクトの前記ディレクトリリスト上のオブジェクトキャッシュのネットワークアドレスを含むメッセージを送信するステップをさらに含む方法。

【請求項12】 請求項10に記載の方法であつて、前記ディレクトリキャッシュから前記オブジェクトの前記ディレクトリリストに掲載されたアドレスを持つオブジェクトキャッシュに、前記オブジェクトキャッシュが前記オブジェクトのコピーを記憶しているかどうかを示すメッセージを受信キャッシュに送信するよう前記オブジェクトキャッシュに要求するメッセージを送信するステップをさらに含む方法。

【請求項13】 請求項10に記載の方法において、複数のディレクトリリストが複数のディレクトリキャッシュ上に記憶されており、各ディレクトリリストは、オブジェクトのネットワークアドレスと、前記オブジェクトのコピーを記憶していると称される少なくとも1つのオブジェクトキャッシュのネットワークアドレスとを含み、前記方法は、

前記ディレクトリリストから α 個のオブジェクトキャッシュのネットワークアドレスを選択するステップと、前記ディレクトリキャッシュから選択された各オブジェクトキャッシュに、オブジェクトキャッシュが前記オブジェクトのコピーを記憶しているかどうかを示すメッセージを前記受信キャッシュに送信するよう要求するメッセージを送信するステップと、

前記選択された各オブジェクトキャッシュから前記受信キャッシュに、前記オブジェクトキャッシュが前記オブジェクトのコピーを記憶しているかどうかを示すメッセージを送信するステップと、

前記オブジェクトのコピーを記憶していることを示す、前記受信キャッシュが受信した最初のメッセージを送信したオブジェクトキャッシュを、主要オブジェクトキャッシュとして特定するステップと、

前記受信キャッシュから前記主要オブジェクトキャッシュに、前記オブジェクトのコピーを前記受信キャッシュに送信するよう要求するメッセージを送信するステップと、をさらに含む方法。

【請求項14】 請求項10に記載の方法において、複数のディレクトリリストが複数のディレクトリキャッシュ上に記憶されており、各ディレクトリリストは、オブジェクトのネットワークアドレスと、前記オブジェクトのコピーを記憶していると称される少なくとも1つのオブジェクトキャッシュのネットワークアドレスとを含み、前記方法は、

前記ディレクトリキャッシュから前記受信キャッシュに、前記オブジェクトのディレクトリリスト上のオブジェクトキャッシュのアドレスを含むメッセージを送信するステップと、

前記ディレクトリリストから α 個のオブジェクトキャッシュのネットワークアドレスを選択するステップと、

前記受信キャッシュから選択された各オブジェクトキャッシュに、そのオブジェクトキャッシュが前記オブジェクトのコピーを記憶しているかどうかを示すメッセージを前記受信キャッシュに送信するよう要求するメッセージを送信するステップと、

前記選択された各オブジェクトキャッシュから前記受信キャッシュに、そのオブジェクトキャッシュが前記オブジェクトのコピーを記憶しているかどうかを示すメッセージを送信するステップと、

10 前記オブジェクトのコピーを記憶していることを示す、前記受信キャッシュが受信した最初のメッセージを送信したオブジェクトキャッシュを、主要オブジェクトキャッシュに特定するステップと、

前記受信キャッシュから前記主要オブジェクトキャッシュに、前記オブジェクトのコピーを前記受信キャッシュに送信するよう要求するメッセージを送信するステップと、をさらに含む方法。

【請求項15】 ネットワーク上のスケーラブルな分散型キャッシングシステムであつて、

20 データオブジェクトのコピーを記憶する手段と、前記オブジェクトのコピーが記憶されているネットワーク位置を特定するディレクトリリストを記憶する手段と、

前記オブジェクトの前記ディレクトリリストが記憶されているネットワーク位置を求める手段と、

前記オブジェクトのコピーを要求する場所を選択する手段と、

前記オブジェクトのコピー入手する手段と、を含むシステム。

30 【請求項16】 請求項15に記載のシステムであつて、オブジェクトが記憶される新しいネットワーク位置を前記オブジェクトの前記ディレクトリリストに加える手段をさらに含むシステム。

【請求項17】 請求項15に記載のシステムであつて、あるオブジェクトの前記ディレクトリリストからネットワーク位置を削除する手段をさらに含むシステム。

【請求項18】 請求項15に記載のシステムであつて、記憶されたオブジェクトのコピーを削除する手段をさらに含むシステム。

40 【請求項19】 請求項15に記載のシステムであつて、あるオブジェクトのディレクトリリストを削除する手段をさらに含むシステム。

【請求項20】 請求項15に記載のシステムであつて、前記オブジェクトのオリジナル情報源からオブジェクトのコピー入手する手段をさらに含むシステム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明はネットワーク上のオブジェクトのキャッシングに関し、より特定的には、スケーラブルな分散型ディレクトリを用いて、キャッシング

ユされたオブジェクトを管理する方法に関する。

【0002】

【背景技術】データオブジェクトのキャッシングは、ネットワーク情報サービスのスケーラビリティの向上に利用される。キャッシングは、大量のデータオブジェクトがネットワークに加えられる場合に発生しうるネットワーク上のサービスの劣化を抑える助けとなる。キャッシングを用いない場合、ユーザがネットワークを通じてデータオブジェクトをリクエストすると、そのデータオブジェクトの情報源にオブジェクトのコピーの送信が要求される。遠隔サイトやデータオブジェクトの数が多い大規模なネットワークでは、この要求を処理するのに長時間かかるか、またはネットワークの状態、ネットワークトラフィック、およびデータオブジェクト情報源の状態によってまったく処理不能となる場合がある。キャッシングを用いるシステムは、オブジェクト情報源以外の最低一箇所、多くの場合はリクエスト元の近くの位置に、データオブジェクトのコピーを記憶する。オブジェクトの別情報源が最低1つ存在すること、およびその別情報源とリクエスト元のユーザとの状況がおそらくはより好都合のために、リクエスト元のユーザにコピーを迅速かつ正しく配信してサービスの向上を計ることができる。

【0003】オブジェクトキャッシングが使用されるネットワークの一例として、インターネットがある。インターネット上でキャッシングを用いるには、オリジナル情報源のサーバ以外にインターネットに接続された最低1つのサーバ上に、(ウェブサイト等からの)ハイパーテキストファイルのコピーをキャッシングする。例えば、日本のサーバ上のある人気のあるファイルを米国のクライアントがリクエストする度に、インターネット上でワールドワイドウェブを介してそのファイルを取り寄せる代わりに、米国内でそのファイルをコピーしてキャッシングサーバに記憶しておくことができる。その後、そのファイルを米国の他のクライアントがリクエストする場合は、日本のサーバではなく米国内のキャッシングサーバからそのファイル入手することができる。この利点は、リクエストおよびコピーのどちらも、リクエスト元と日本のオリジナル情報源のサーバとの間よりもキャッシングサーバとの間のほうが、インターネット中で通過するルータおよびノード(中継点)の数が少なくてすむことである。このため、コピー入手までの遅れが少なくなり、かつリクエスト元のユーザがコピーを正しく受信できる可能性が高くなる。こうして、米国のクライアントが日本のファイル入手する費用および待ち時間(遅れ)を大幅に低減することができる。

【0004】キャッシングを採用するシステムは、3つの機能上の構成要素からなることが理解できる。第1に、キャッシングは、ユーザによってリクエストされたデータオブジェクトのキャッシングされたコピーの場所を決定する検索システムを含む。第2に、キャッシングを採

用するシステムは、リクエストされたオブジェクトの1つ以上のキャッシングされたコピーの場所を検索システムが突き止めた場合に、どのコピーを取り出すのかを決定する決定システムを含む。最後に、維持システムは、古い期限切れのオブジェクトのコピーをキャッシングから取り除き、キャッシングによって記憶された情報の正確さを全体的に維持する助けをする。

【0005】図1に従来のキャッシングシステムの階層構造を示す。クライアント101はキャッシングサーバA

10 102にデータオブジェクトをリクエストする。以下、「キャッシング」はキャッシングサーバをさるものとする。キャッシングA102(クライアントからオブジェクトのリクエストを受信するキャッシング)は、一般にはネットワーク的にクライアントの「近くに」位置する。つまり、キャッシングAとクライアントとは、直接またはネットワークを介して迅速かつ信頼性をもって、または少なくともクライアント101がキャッシングA以外のサーバと通信するよりは迅速かつ信頼性をもって、互いに通信しあえる。ネットワーク的にクライアントの「近くに」位置するキャッシングサーバは、ネットワーク的に「近くない」他のサーバよりも物理的には離れている場合があることを理解されたい。

【0006】クライアントのリクエストは、クライアント101とキャッシングA102とを接続するネットワーク(図示せず)を介して送信できる。もしキャッシングAがオブジェクトのコピーを持っていれば、キャッシングAはそれをクライアント101に送信する。もしリクエストされたデータを持っていなければ、キャッシングA102はキャッシングされたコピーを検索しなければならない。この場合、キャッシングA102はまず、階層レベル

30 が同じ近隣キャッシング群、つまりキャッシングB103、キャッシングC104、およびキャッシングD105にリクエストを送信する。キャッシングA102、キャッシングB103、キャッシングC104、およびキャッシングD105は、ネットワーク(図示せず)で接続できる。各キャッシングは、応答中のキャッシングがリクエストされたオブジェクトのコピーを持っているかどうかを知らせるメッセージをキャッシングA102に送って、キャッシングA102からのリクエストに応じる。同階層レベルのキャッシング群がリクエストされたオブジェクトを持っていなければ、キャッシングA102は次の階層レベルのキャッシングE106にオブジェクトのリクエストを送信する。キャッシングA102は、ネットワーク(図示せず)によってキャッシングE106に接続できる。同様に、もしキャッシングE106がオブジェクトを持っていなければ、キャッシングE106は同じ階層レベルのキャッシング群、つまりキャッシングF107およびキャッシングG108をポーリングする。キャッシングE106、キャッシングF107、およびキャッシングG108はネットワーク(図示せず)で接続できる。これらのキャッシング群のどれもオブ

40 ジェクトを持っていなければ、キャッシングE106は同じ階層レベルのキャッシング群、つまりキャッシングF107およびキャッシングG108をポーリングする。キャッシングE106、キャッシングF107、およびキャッシングG108はネットワーク(図示せず)で接続できる。これらのキャッシング群のどれもオブ

ジェクトを持っていない場合は、オブジェクトのオリジナルバージョンが存在するサーバ109にメッセージが送られる。サーバ109に送信されるメッセージは、リクエストされたオブジェクトのコピーをクライアント101に送信するようサーバ109に要求する。いずれか1つのキャッシングがオブジェクトのコピーを持っているというメッセージを送信する場合は、そのキャッシングがクライアント101にオブジェクトのコピーを送信するよう要求される。一方、1つ以上のキャッシング群が、オブジェクトのコピーを持っているというメッセージを送信する場合は、決定関数を実行して、どのキャッシングがオブジェクトのコピーをクライアント101に送信するかを選択しなければならない。

【0007】図2～図4は公知の決定システムを示す。クライアント201はキャッシングA202にオブジェクトのリクエスト(図2では“request(1)”で示す)を送信する。キャッシングA202はコピーを持っているかどうかを判断する。もし持っていないれば、キャッシングA202はクライアント201にコピーを送信する。もし持っていないければ、キャッシングA202はすべての近隣キャッシング群203～207に、“UDP_ping_req(2)”で示すUDP_pingリクエストと呼ばれるリクエストを送信する。各メッセージの後の括弧内の数字(“request(1)”、“UDP_ping_req(2)”等)は、メッセージが送信される順番を示す。

【0008】図3に示すように、各近隣キャッシングは、もしオブジェクトのコピーを持っていればUDP_pingヒットメッセージ(図3では“UDP_ping_hit”で示す)で応答し、もし持っていないればUDP_pingミスマッチメッセージ(図3では“UDP_ping_miss”で示す)で応答する。UDP_pingヒットメッセージがまったく受信されない場合は、リクエストされたオブジェクトのコピーの送信はオリジナル情報源のサーバに依頼される。キャッシングA302がUDP_pingヒットメッセージを1つだけ受信する場合は、キャッシングA302はコピーを持っている唯一のキャッシングにそのオブジェクトのコピーをクライアント301に送信するよう要求するメッセージを送信する。図3に示すように1つ以上のUDP_pingヒットメッセージを受信する場合は、キャッシングA302は、どのキャッシングにオブジェクトコピーをクライアント301へ送信するよう依頼すべきかを決定しなければならない。この先行技術では、図4に示すように、キャッシングA402は、UDP_pingヒットメッセージを最初に受信したキャッシング、つまりキャッシングE403(図4)に、オブジェクトのコピーを要求するメッセージ(“REQUEST(1)”)を送信する。この決定関数は、ネットワーク的にキャッシングA302に(従ってクライアント301に)「最も近い」コピーを入手

するように設計されている。このため、リクエストされたオブジェクトのコピーを、最小の遅れと最大の信頼性で入手することができる。

【0009】コピーはキャッシングE403からキャッシングA402へ送られ(COPY(2))、ネットワーク的にはキャッシングEよりもクライアント401に「より近い」キャッシングA402にオブジェクトのコピーが記憶される。この結果、このオブジェクトに対するクライアント401からのこれ以降のリクエストでは、キャッシングA402以外の検索は不要となる。次に、オブジェクトのコピーはキャッシングA402からクライアント401へ送信される(COPY(3))。先行技術の他の実施形態では、リクエストは、キャッシングA402をバイパスして、キャッシングE403にオブジェクトのコピーを直接クライアント401に送ることを要求する。

【0010】キャッシングを採用したシステムのこの先行技術の実施形態は、待ち時間(リクエストとオブジェクトのコピーの受信との間の遅れ)を短縮し、比較的データオブジェクトの数が少ない小規模なネットワーク中では、リクエストされたオブジェクトのコピー入手の信頼性を向上させる。しかし大規模なネットワークでは、かかる公知の方法はスケーラビリティ(拡張性)が低い。例えばオブジェクトに対するクライアントからのリクエストを受信するキャッシング(以後、「受信キャッシング」と称する)がN個の近隣キャッシングを持つキャッシングのシステムを考える。これらのキャッシング群は、リクエストされたオブジェクトのコピーを保有するメッセージ以外に、各クライアントリクエストごとに2N+1個のメッセージを作成する。つまり、各近隣キャッシングごとに2つのメッセージ(UDP_pingリクエストと、UDP_pingヒット応答またはUDP_pingミスマッチ応答のいずれか)と、クライアントからのリクエストメッセージとである。このように大量のメッセージは、データオブジェクト数の多い大規模ネットワークでは大変な負担となる。1つのキャッシングが追加されるごとに、キャッシングシステムが作成するメッセージ数が2つずつ増加する。

【0011】この問題を解決するための公知の方法は、ディレクトリ(例えば、データオブジェクトを識別する情報と、そのキャッシング場所のリスト)を使用して、データオブジェクトのIDを該オブジェクトが記憶されているキャッシング場所に対応づけることである。これにより全キャッシングをフルに検索する必要がなくなり、キャッシングシステムがネットワークに課すメッセージトラフィック量を少なくできる。例えば、メッセージをすべての近隣キャッシング群へ送信する代わりに、1つのリクエストをディレクトリをもつ中央キャッシングへ送信して、中央キャッシングにリクエスト元に一番近い、リクエストされたコピーの場所を問い合わせせる。

【0012】ディレクトリを使用すると、キャッシングさ

れたオブジェクトのコピーの場所をつきとめて入手するのに必要なトラフィック量を低減できるが、キャッシングコヒーレンスの問題が生じる。つまり、データオブジェクトが変更されると、維持システムによってそのデータオブジェクトの全コピーを変更し、かつ全ディレクトリを更新しなければならない。分散型コヒーレンスプロトコルを用いてオブジェクトとディレクトリとのコヒーレンスを維持し、ネットワークを介して送信されるメッセージを用いて古いオブジェクトのコピーを無効にし、ディレクトリのエントリを更新する。しかし、このメッセージトラフィックはネットワーク、特に変更頻度の高い多数のオブジェクトをもつ大規模ネットワークに負担をかけてしまう。

【0013】オブジェクトのコピーが古くなる（つまりオリジナルのデータオブジェクトが変更される）と、無効コピーは無効化されてキャッシング群から除去または排除される。さらに、キャッシングされたコピーをさすディレクトリが更新される。これは従来のいくつかのシステムでは、各コピーに生存時間（TTL）パラメータを割り当てることによって行われる。TTLパラメータは、記憶されたデータオブジェクトのコピーが期限切れして削除されるべき日時を特定する。TTLが到来すると、コピーはキャッシングシステムから排除される（つまりキャッシングから削除される）。TTLが期限切れする前に古くなったコピーを排除しようとすれば、従来の大規模なキャッシングシステム中ではトラフィックの負担が大きすぎて実現できない。従って、従来のシステムはTTLパラメータに基づいて弱いコヒーレンスしか維持できない。コヒーレンスが弱いのは、キャッシング化されたからすぐに変更されたオブジェクトのキャッシングされたコピーが、TTLが切れるまで、古くなってから長時間にわたって利用可能なまま残留するためである。この問題は、オブジェクトのコピーおよびディレクトリ中のそのエントリを削除するメッセージをばらまくことによって解決できるが、これによりトラフィックでネットワークにさらに負担をかけてしまう。

【0014】ロッキングおよび承認も、従来のキャッシングシステムのスケーラビリティを下げる原因である。従来のシステム中でディレクトリが更新される場合、変更中はディレクトリがロック（つまりアクセス阻害）されなければならない。このため更新中はディレクトリは利用不可となる。また、更新の受信を示すために承認メッセージが送信されなければならず、ネットワークの一層の負担となる。従って、ロッキングまたは承認が不要なスケーラブルなディレクトリベースのキャッシングシステムであれば、従来のキャッシングシステムよりもネットワークの性能をより効率的かつ効果的に改善することができる。

【0015】

【課題を解決するための手段】本発明は、ネットワーク

上でロッキングおよび承認なしでオブジェクトキャッシングを有利かつ効率的に管理する、スケーラブルな分配型ディレクトリベースのキャッシング方式システムを提供する。このシステムは、キャッシングネットワークトラフィックが少なく、メモリおよび処理電力が従来のキャッシングシステムより少なくてすむ。

【0016】本発明に従えば、オブジェクトに対するユーザリクエストは受信キャッシングによって受信される。受信キャッシングは、リクエストされたオブジェクトのネットワークアドレスを入力として用い、出力としてポインタを与えるロケータ関数を実行する。一実施形態では、ポインタは、オブジェクトのディレクトリリストを記憶するディレクトリキャッシングのネットワークアドレスを指す。オブジェクトのディレクトリリストは、オブジェクトのネットワークアドレスと、オブジェクトのコピーを記憶していると称されるオブジェクトキャッシング群のネットワークアドレスとを含む。

【0017】受信キャッシングはディレクトリキャッシングにオブジェクトリクエストメッセージを送信し、ディレクトリキャッシングはディレクトリリストの掲載されたオブジェクトキャッシング群をポーリングする。これに応答して、オブジェクトキャッシング群は、各オブジェクトキャッシングがオブジェクトのコピーを記憶しているかどうかを示すメッセージを受信キャッシングに送信する。

【0018】受信キャッシングは、オブジェクトのコピーを記憶していることを示すメッセージであって、受信キャッシングが受信した最初のものを送信したオブジェクトキャッシングに、オブジェクトリクエストメッセージを送信する。これに応答して、送信元のオブジェクトキャッシングは、受信キャッシングにオブジェクトのコピーを送信し、受信キャッシングはコピーを記憶して、ユーザにコピーを転送する。

【0019】ディレクトリリストはその後、受信キャッシングのネットワークアドレスを追加して更新される。古くなったりキャッシングされたオブジェクトのコピーは、オブジェクトのオリジナル情報源のサーバからキャッシングに新しいコピーが送信されるごとに、分散方式で削除される。キャッシングされたコピーのコヒーレンスは、生存時間（time-to-live）パラメータを各コピーに関連づけることによってさらに向上できる。生存時間パラメータはまた、オブジェクトのディレクトリリスト上のオブジェクトキャッシングのアドレスに関連づけられる。

【0020】本発明の分散型機能によって、本発明はスケーラブルかつ効率的となり、データオブジェクトのコピーをキャッシングし、キャッシングされたコピーを迅速かつ経済的にユーザに与えることができる。

【0021】

【発明の実施の形態】本発明は、ネットワーク上でキャッシング群を用いて実現されるスケーラブル分散型キャッシ

11

シングルシステムを提供する。キャッシュはサーバであり、プロセッサとコンピュータにより読み出し可能なメモリとネットワークに接続されるポートとの間を接続するデータバスを含む。本発明に従うキャッシュは、データオブジェクトのコピーを記憶する。キャッシュはまた、オブジェクトのコピーを記憶するキャッシュ群のアドレスを掲載するオブジェクトのディレクトリリストを記憶（または該リストのみを記憶）する。キャッシュはまた、オブジェクトのネットワークアドレスに基づいて、オブジェクトのディレクトリリストを記憶するキャッシュのネットワークアドレスの位置を求める際に有用なハッシュロケータ関数を実現できる。本明細書において、「オブジェクトのネットワークアドレス」はオリジナルオブジェクトの情報源のネットワークアドレス上の位置を特定可能なあらゆるものをさす。オブジェクトのネットワークアドレスの一例は、オブジェクトのオリジナル情報源におけるユニフォーム・リソース・ロケータ（URL）である。「ネットワークアドレス」をサーバ*

Directory_list

```

{
    directory_object_address
        /* リストオブジェクトのネットワークアドレス */
    {
        /* リストオブジェクトのコピーが記憶されている
           5つのオブジェクトキャッシュ群のアドレス */
        object_cache_address_1
        object_cache_address_2
        object_cache_address_3
        object_cache_address_4
        object_cache_address_5
    }
}

```

クライアントからオブジェクトリクエストを受信する各キャッシュは、ハッシュロケータ関数を備える。ハッシュロケータ関数の入力はオブジェクトのネットワークアドレスであり、出力はリクエストされたオブジェクトのディレクトリリストが記憶されているディレクトリキャッシュのアドレス（またはそのポインタ）である。こうして、受信キャッシュはキャッシュされた任意のオブジェクトについてディレクトリキャッシュの位置を求めることができる。リクエストされたオブジェクトにディレクトリリストがない（つまりロケータ関数がどのディレクトリキャッシュも指さない）場合は、リクエストはオリジナル情報源のサーバに送られ、リクエストされたオブジェクトを受信キャッシュに送信させる。受信キャッシュはリクエストされたオブジェクトのコピーを記憶し、そのコピーをクライアントに送信する。その後、受信キャッシュは、オブジェクトキャッシュとして受信キャッシュアドレスを含むオブジェクトのディレクトリリストを開始させる。一実施形態では、受信キャッシュは

50

【0025】

12

*（キャッシュ等）に関して用いる際は、ネットワーク上でサーバ位置を特定可能なあらゆるものとします。「ネットワークアドレス」のあるオブジェクトのキャッシュされたコピーに関して用いる際は、ネットワーク上でオブジェクトのキャッシュされたコピーの位置を求めることが可能なあらゆるものとします。オブジェクトのキャッシュされたコピーのネットワークアドレスの一例はURLである。

【0022】本発明に従えば、各キャッシュされたデータオブジェクトはディレクトリリストを有する。オブジェクトのディレクトリリスト上の各アドレスは、オブジェクトのコピーを記憶するキャッシュ（オブジェクトキャッシュ）のアドレスである。好適な実施形態では、ディレクトリリスト群は複数のキャッシュ（ディレクトリキャッシュ）間に分散される。本発明に従うディレクトリリストのデータ構造の一実施例を以下に示す。

【0023】

【表1】

オブジェクトのディレクトリリストも記憶する。この例からわかるように、キャッシュは、受信キャッシュ（クライアントのリクエストを受信）、ディレクトリキャッシュ（オブジェクトのディレクトリリストを記憶）、およびオブジェクトキャッシュ（オブジェクトのコピーを記憶）の機能を同時に行うことができる。

【0024】図5は、ネットワークアドレスYをもつデータオブジェクトへのリクエスト（図5では“GET_request(Y)”で示す）をクライアント502から受信する受信キャッシュであるキャッシュA501を示す。受信キャッシュA501は、アドレスYを入力として用いてハッシュロケータ関数Lを実行し、ディレクトリキャッシュアドレスCを出力として得る。これはつまり、キャッシュC503がオブジェクトYのディレクトリリストを記憶していることを示す。オブジェクトYのディレクトリリストのデータ構造の一実施例を以下に示す。

【表2】

Directory_list

```

{
    Y /* オブジェクトYのネットワークアドレス */
{
/* それぞれディレクトリリストオブジェクトYの
   コピーを記憶しているオブジェクトキャッシュ
   B, D, E, およびF のアドレス */
B
D
E
F
}
}

```

上記のデータ構造は、データオブジェクトYのコピーがオブジェクトキャッシュB 504, D 505, E 506, およびF 507に記憶されていることを示す。受信キャッシュA 601(図6)は、データオブジェクトをリクエストするメッセージ(図6では“LOOKUP”で示す)をディレクトリキャッシュC 602に送信する。図6に示す実施形態では、ディレクトリキャッシュC 602は受信キャッシュA 601に、ディレクトリリストからのオブジェクトキャッシュアドレス群を含むメッセージ(図6では“LIST [B, D, E, F]”で示す)を送信する。キャッシュA 601は、このリストを用いてオブジェクトキャッシュ群をポーリングし、その中からオブジェクトYを受信するキャッシュを1つ選択する。本発明の他の実施形態では、ディレクトリキャッシュC 602はオブジェクトキャッシュのアドレス群を含むメッセージを受信キャッシュA 601に送信せず、代わりにディレクトリキャッシュC 602自身がオブジェクトキャッシュ群(図6の実施形態ではB-603, D-604, E-605, およびF-606で示す)をポーリングし、各オブジェクトキャッシュにUDP pingヒットまたはミスマッチメッセージを受信キャッシュA 601に送信するよう依頼する。その後、受信キャッシュA 601は、どのオブジェクトキャッシュからオブジェクトYのコピーを要求するかを決定し、選択したオブジェクトキャッシュにコピーの送信を要求するメッセージを送る。

【0026】本発明は、所与のオブジェクトのキャッシュされたコピーの位置を求めるために非分散型ディレクトリを使用する従来のディレクトリシステムよりも効率的である。ある従来のシステムでは、キャッシュされたオブジェクトに対するリクエストがあるたびに、1つのサーバに存在する1つの非分散型ディレクトリに問い合わせが行われる。この単独のディレクトリは、リクエストされたオブジェクトが記憶されているキャッシュの最低1つのアドレスを返送する。この従来のシステムの欠点は、単独サーバがキャッシュされたオブジェクト群に対する全リクエストのボトルネックとなり、リクエスト

されたオブジェクトを迅速かつ信頼性をもって返信するシステムの性能の向上というキャッシングの目的を妨げることである。クライアントと単独ディレクトリサーバとの間のリクエストメッセージおよび応答メッセージのトラフィックによって、またディレクトリの正確さを維持するための維持トラフィックによって、ネットワークに過度の負担がかけられる。単独ディレクトリサーバに対する負担が重くなっている場合がある。さらに、非分散型サーバは、サーバ上でかなり大量のメモリリソースをとるためにサイズが大きくなり、探索速度が落ちてプロセッサの処理時間が長くなる場合がある。また単独ディレクトリサーバは、キャッシングシステムの単独故障点でもある。もし単独ディレクトリサーバが故障すれば、キャッシングシステム全体が動作不能になり、キャッシングされたコピーをオブジェクトキャッシュ群に記憶するのに使用されたすべてのメモリ資源が無駄になってしまう。

【0027】従来のシステムの問題の一部を解決する明らかな方法は、ネットワーク上の複数の異なる場所(すなわちネットワークアドレスの異なる複数の異なるサーバ上)に、非分散型ディレクトリの複数のコピーを置くことである。これにより、クライアントはオブジェクトキャッシュのアドレスを入手するのに、多数の非分散型ディレクトリの1つに問い合わせができるから、単独ディレクトリサーバを有することによって生じるボトルネックおよび単独故障点の影響を減じることができる。これはディレクトリ探索活動の負担を分散させる助けとなる。また、1つのディレクトリサーバが故障しても、別のサーバに問い合わせを行うことができる。しかし、この解決方法では、ディレクトリ群の正確さを維持する必要があるため、ネットワーク上のキャッシングシステムトラフィックによる負担を悪化させてしまう。従って、キャッシングされたコピーの状態が変更される(キャッシングされたコピーが古くなったり、新たなオブジェクトがキャッシングされる等)と、全ディレクトリサーバに更新メッセージの送信が要求されるため、承認メッセージ

の形式でさらに大量のトラフィックを発生させてしまう。これにより、かなり大量のメッセージトラフィックを発生させてネットワークに負担をかけてしまうという欠点がある。また、単独非分散型ディレクトリの場合と同じく、従来のシステムにおける1つ以上のディレクトリをもつ非分散型ディレクトリ群は、更新中はロックされなければならず、その間は使用できない。オブジェクトの状態が頻繁に変更される大規模なシステムでは、更新およびロッキングは非常に頻繁に発生するため、探索ディレクトリの利用可能性を大幅に下げてしまう。また、大規模な非分散型ディレクトリの複数のコピーをもつことにより、メモリおよびプロセッサ資源を奪ってしまうという問題が悪化する可能性がある。

【0028】本発明はこのような問題を解決し、従来のシステムより効率よくかつ経済的に使用および維持可能な、キャッシングシステム用の分散型ディレクトリを提供する。本発明に従ってディレクトリをディレクトリリスト群に分割して、これらのリストを多数のサーバに分散させることによって、従来のシステムよりも多数のサーバにディレクトリ参照の負担を分散することができる。これによりボトルネックおよび単独故障点の問題がかなり低減または解消される。非分散型ディレクトリキャッシングシステムのディレクトリサーバと比べて、各ディレクトリサーバ上のディレクトリの記憶に必要なメモリ資源が大幅に少なくなるため、サーバはディレクトリデータ以外のデータを記憶できる。さらに、このようにディレクトリ探索の負担が有利に分散されるため、任意の1つのサーバが吸収するプロセッサ時間がはるかに少なく、ディレクトリサーバが探索リクエストに迅速かつ効率的に応答できる。本発明の他の利点は、ディレクトリ維持に必要なトラフィックレベルが低く、従来のシステムに比べてネットワーク上の負担が少ないことである。本発明は承認メッセージを使用せず、有利には更新時にディレクトリデータをロックせず、かつ承認メッセージを使用しない。

【0029】検索システムは、リクエストされたオブジェクトのコピーが記憶されている場所がもしあれば、その場所を求める。上述したように、本発明に従えば、キャッシングはクライアントからオブジェクトの入手を求めるリクエスト（“GET request”と称する）を受信する。図5は、クライアント502が入手リクエストをキャッシングシステムA501に送信するキャッシングシステムの一実施形態を示す。リクエストを受信すると、受信キャッシングであるキャッシングA501はまず、要求されたオブジェクトを自身が記憶しているかどうかを判断する。キャッシングA501がリクエストされたオブジェクトを持っている場合は、該オブジェクトがクライアント502に送信される。持っていない場合は、キャッシングA501はどこか他の場所でオブジェクトの位置を求めなければならない。

【0030】リクエストされたオブジェクトを持っていない場合、キャッシングA501はディレクトリリストロケータ関数L508を実行する。上述したように、ディレクトリリストロケータ関数L508は、リクエストされたデータオブジェクトのネットワークアドレスを入力として使用して、該オブジェクトのディレクトリリストを記憶するキャッシングのアドレス（またはそのポインタ）を出力する。このため、キャッシングA501がクライアント502からデータオブジェクトYに対するリクエストを受信し、キャッシングA501がオブジェクトYを記憶していないと判断すると、キャッシングAはL508を実行して、Yのディレクトリリストの位置を求める。図5に示すL(Y)=Cは、キャッシングC503がオブジェクトYのディレクトリリストを記憶していることを示す。

【0031】本発明のある実施形態においては、ロケータ関数はネットワークアドレスを示さない。これは、ロケータ関数を実行する受信キャッシングの論理上、到達可能な範囲内のキャッシングにオブジェクトのコピーがない、またはどのキャッシングにもオブジェクトのコピーが記憶されていないことを意味する。受信キャッシングの論理範囲内のキャッシングにオブジェクトのコピーが記憶されていない本発明の実施形態では、キャッシング群は論理レベルの階層中で論理レベルごとにグループ分けされる。第1の論理レベルのキャッシング群の1つは、同レベルの他のキャッシングに記憶されているオブジェクトのコピーを検索して入手できる。第2レベルのキャッシングに記憶されているオブジェクトのコピー入手するには、第2レベルの受信キャッシングにオブジェクトリクエストメッセージが送信され、第2レベルの受信キャッシングは第2レベルのキャッシングからオブジェクトのコピーの検索を開始し、該コピーの入手を試みる。コピー入手すると、そのコピーは第1レベルの受信キャッシングへ送られる。本実施形態では、ロケータ関数が、第1レベルの受信キャッシングが論理上利用可能なキャッシングのどれもがオブジェクトのディレクトリリストを記憶していないことを示すポインタを返送する場合は、第1レベルの受信キャッシングは第2レベルの受信キャッシングへのオブジェクトリクエストの送信を選択する。

【0032】また、本発明の他の実施形態では、ロケータ関数から返送されたポインタが、受信キャッシングが論理上利用可能なキャッシングのどれもがオブジェクトのディレクトリリストを記憶していないことを示す場合は、受信キャッシングはオブジェクトのオリジナルバージョンを記憶しているサーバにオブジェクトのコピーをリクエストすることを選択する。

【0033】このように、ハッシュ関数L508は、オブジェクトのネットワークアドレスを入力として使用し、出力としてポインタを与えて、該ポインタが示す情報が本発明のこれ以降の行動を決定する関数を含む。ボ

インタがオブジェクトのディレクトリリストを記憶しているディレクトリキャッシュのネットワークアドレスを示すと、オブジェクトリクエストは受信キャッシュからディレクトリキャッシュへ送られる。一方、ポインタがネットワークアドレスをまったく示さない場合、または論理的に利用可能なキャッシュのネットワークアドレス（同じ階層レベルのキャッシュのアドレス等）を示さない場合は、オブジェクトリクエストメッセージは他の階層レベルの受信キャッシュに送られるか、またはオブジェクトのオリジナル情報源サーバへ送られる。一実施形態では、ロケータ関数のポインタをディレクトリキャッシュのネットワークアドレス等の情報と相關させるテーブルが、システム管理者によって手動で受信キャッシュに入力される。他の実施形態では、ポインターテーブルを更新する受信キャッシュにメッセージが送られる。例えば、新たにキャッシュされたオブジェクトXについて新たなディレクトリリストがディレクトリキャッシュZにおいて開始されると、オブジェクトXのネットワークアドレス、ディレクトリキャッシュZのネットワークアドレス、およびポインターテーブルが更新されるべきであることを示すフラグを含むメッセージが、受信キャッシュに送信される。受信キャッシュはこのメッセージを受信し、オブジェクトのネットワークアドレスに対してハッシュ関数を実行し、得られたポインタ値をキャッシュZのネットワークアドレスに相關させる。このように、オブジェクトXに対するリクエストを受信すると、受信キャッシュはXに対してハッシュロケータ関数を行い、ポインタを入手し、そのポインタをキャッシュZのネットワークアドレスに相關させる。その後、受信キャッシュは、本発明に従ってキャッシュZに探索リクエストを送る。

【0034】図5に示す例では、キャッシュA501はオブジェクトYの入手リクエストを受信している。キャッシュA501は、それ自身はオブジェクトYを記憶していないと判断し、Yのネットワークアドレスに対してロケータ関数L508を実行している。本例では、 $L(Y) = C$ であり、キャッシュC503がオブジェクトYのディレクトリリストを記憶していることを示す。

【0035】図6は、図5においてディレクトリリストロケータ関数L508が、キャッシュCがオブジェクトYのディレクトリキャッシュを記憶していると特定した後の様子を示す図である。キャッシュA601はキャッシュC602に“探索リクエスト（LOOKUP request）”を送信する。一実施形態では、探索リクエストは、リクエストされたオブジェクトYのアドレスと、リクエスト元のキャッシュつまりキャッシュA601のアドレスとを含む。キャッシュC602は、リクエストされたオブジェクトYのディレクトリリストキャッシュのアドレス群をキャッシュA601に返送する。キャッシュA601は決定システムを実行して、この中か

らオブジェクトのコピーを要求するキャッシュを選択する。他の実施形態では、キャッシュC602はキャッシュA601にディレクトリリストを送信せず、この代わりにキャッシュC602は、リクエストされたオブジェクトのディレクトリリスト上のオブジェクトキャッシュ群に、キャッシュA601のネットワークアドレスとともに、UDP pingリクエストを送信する。オブジェクトキャッシュ群は、キャッシュA601にUDP pingヒットメッセージまたはミスマッチングで応答する。

【0036】本発明の主たる利点はそのスケーラビリティである。ディレクトリがディレクトリリストの形で多数のサーバ間に分配されるため、十分なスケーラビリティをもつ検索機能を実現できる。所与のオブジェクトに対する正しいディレクトリリストは受信キャッシュ中で実行されるハッシュロケータ機能によって見つけだされる。このように、多数のリクエストによる負担を受け、かつ1つのサーバ上に存在する単独の非分散型ディレクトリリストを完全に検索する代わりに、ハッシュロケータ関数と短いディレクトリリストとを参照することによって、ネットワーク上の多数の異なるサーバ上にキャッシュされた多数の異なるオブジェクトを簡単かつ効率的に発見できる。

【0037】リクエストされたオブジェクトのコピーをもつキャッシュ群のアドレスが判明すると、決定システムが実行されて、リクエストされたオブジェクトのコピーをどのキャッシュから送信するかを決定する。一実施形態では、リクエストされたオブジェクトのキャッシュされたコピーがなければ、受信キャッシュはオブジェクトに対するリクエストをオブジェクトのオリジナル情報源のサーバに送信する。他の実施形態では、受信キャッシュはオブジェクトに対するリクエストを、キャッシングシステムの他の階層レベルの別のキャッシュに送信する。

【0038】上述したように、本発明の一実施形態では、受信キャッシュからの探索リクエストに応答して、ディレクトリキャッシュから受信キャッシュへオブジェクトキャッシュのアドレス群が送られ、この場合は受信キャッシュがオブジェクトキャッシュ群をポーリングする。他の実施形態では、オブジェクトキャッシュアドレス群は受信キャッシュへは送られず、この場合はディレクトリキャッシュがオブジェクトキャッシュ群をポーリングする。いずれの場合も、オブジェクトキャッシュ群から受信キャッシュへUDP pingヒットまたはミスマッチングが送られ、受信キャッシュはリクエストされたオブジェクトのキャッシュされたコピーの送信をどのオブジェクトキャッシュに依頼するかを決定する。

【0039】図7は、オブジェクトキャッシュアドレス群が受信キャッシュに送信されない場合の実施形態を示す。キャッシュA701はクライアントからオブジェク

トYに対するリクエストを受信している。キャッシュA701はディレクトリ探索関数 $L(Y) = C$ （図示せず）を実行し、オブジェクトYのディレクトリリストがキャッシュC702に記憶されていることを示す。受信キャッシュA701はディレクトリキャッシュC702に探索メッセージ（“LOOKUP(1)”）を送信する。この例では、キャッシュC702に記憶されているYのディレクトリリストは、YのコピーがキャッシュE703とF704とに記憶されていることを示す。ディレクトリキャッシュC702は、オブジェクトキャッシュE703とF704とにUDP pingリクエストを送信する。図8に示すように、キャッシュA801はまず、キャッシュE802からUDP pingヒット応答（“UDP_ping_hit(1)”）を受信し、その後、キャッシュF803からUDP pingヒット応答（“UDP_ping_hit(2)”）を受信する。図9に示すように、キャッシュA901が受信した最初のUDP pingヒット応答はキャッシュE902からであったので、キャッシュA901はキャッシュE902にオブジェクトYのコピーのリクエストを送信する。この決定規準の有利な点は、キャッシュA901がオブジェクトキャッシュ群をポーリングしてからキャッシュE902からUDP pingヒットメッセージを受信するまでの応答時間のほうが、キャッシュF803（図8）からUDP pingヒットメッセージを受信するまでの応答時間よりも遅延が少ないことに基づいて、ネットワーク的に受信キャッシュA901に「一番近い」キャッシュにオブジェクトがリクエストされることである。キャッシュA901からのリクエストに応答して、キャッシュE902はキャッシュA901にオブジェクトYのコピーを送信し、キャッシュA901はそのコピーを記憶してリクエスト元のクライアントへ転送する。

【0040】この時点で、オブジェクトYのディレクトリリストはキャッシュA901がオブジェクトYのコピーを持っていることを反映していないため、不正確となる。本発明の一実施形態では、キャッシュA901はキャッシュC903にディレクトリリストを更新するようメッセージを送る。キャッシュC903はオブジェクトYのディレクトリリストにキャッシュA901のネットワークアドレスを加える。キャッシュA901はこうしてYのオブジェクトキャッシュとなる。

【0041】本発明の他の実施形態では、キャッシュC702（図7）は、リクエストされたオブジェクトのディレクトリリストにアドレスが掲載されているすべてのキャッシュにはUDP pingリクエストを送信しない。この代わりに、キャッシュC702はリストから一定数 α のアドレス群を選択する。図7に示す例においてオブジェクトYのディレクトリリストはB705, D706, E703, およびF704であるとする。キャッシュ

シユC702は α を2に選択し、キャッシュA801のアドレスとともにUDP pingリクエストを送信するキャッシュとしてディレクトリリストからキャッシュE703およびF704を選択する。キャッシュE802（図8）のUDP pingヒットメッセージが最初にキャッシュA801によって受信されるので、キャッシュA801はオブジェクトのリクエスト（“request(1)”（図9））をキャッシュEに送る。リクエストされたオブジェクトのコピー（“COPY(2)”）はキャッシュE902からキャッシュA901（図9）に送られる。オブジェクトのコピーはキャッシュA901に記憶され、コピー（“COPY(3)”）がリクエスト元のクライアント904に送信される。キャッシュA901は、YのディレクトリリストにキャッシュA901のアドレスを加えて更新するようキャッシュC903にメッセージを送る。

【0042】 α がディレクトリリストに掲載されたアドレス数より少ないとときは常に、キャッシュは、ディレクトリリストのどのキャッシュにUDP pingリクエストを送信するかを決定しなければならない。一実施形態では、有利には、キャッシュはラウンドロビン方式（スケジューリング）を用いて、リストに掲載されたキャッシュ間に負担を分散できる。他の実施形態では、キャッシュはリクエストを送るキャッシュ数 α をランダムに選択する。

【0043】本発明の一実施形態では、ディレクトリキャッシュはUDP pingリクエストを送るキャッシュをディレクトリリストから1つ、つまり $\alpha = 1$ に選択する。例えば、キャッシュC1001（図10）がキャッシュE1002を選択するとする。キャッシュC1001はキャッシュA1003のアドレスとともにUDP pingリクエストメッセージをキャッシュE1002に送り、キャッシュE1002は直接キャッシュA1003に応答（UDP pingヒットまたはミス）を送信する。もし応答がUDP pingヒットならば、キャッシュA1003はオブジェクトのリクエストをキャッシュE1002に送る。もし応答がUDP pingミスならば、ディレクトリリストは正しくないから、キャッシュA1101（図11）はキャッシュC1102にディレクトリ削除リクエストを送り、キャッシュC1102はリクエストされたオブジェクトのディレクトリリストからキャッシュE1103のアドレスを削除する。一方、キャッシュEがキャッシュAからのオブジェクトリクエストに応答しない場合は、キャッシュE、またはキャッシュEへの接続が作動不能かもしれないなどで、リスト上の別のキャッシュを選択してUDP pingリクエストを送信しなければならない。

【0044】 $\alpha = 1$ の場合、コピーの検索に必要なメッセージの数は、リクエストされたオブジェクトのコピーをキャッシュC1001からキャッシュA1003に運

ぶメッセージを除き、有利なことに4つに減らされる。これは、同じ目的のために $2N+1$ 個のメッセージが必要な、図2～図4に示す従来の決定関数よりはるかに効率的である。本発明に従って $\alpha = 1$ について決定関数を実行することにより、接続または装置の故障の確率が非常に低く、かつキャッシュC1001が全キャッシュ間に負担を分散する効率的な方法を実行する信頼性のあるネットワークでは、良好なキャッシュ性能が得られる。決定システムが実行されるとディレクトリリストが更新される方式は、信頼性の高いネットワーク（信頼性の高い装置および接続をもつネットワーク等）について有効に動作するため、「楽観的アプローチ」と呼ばれる方式を示す。

【0045】ネットワークの中には信頼性が完全ではないものがある。かかるネットワークの一例はインターネットであり、所与の装置または接続が動作しているかどうか、またはある任意の時間に利用不能かどうかの予測が困難または不可能である。さらに、ディレクトリキャッシュと受信キャッシュとがデータオブジェクトに対するリクエストをどこに送るかについてそれぞれ独自に決定を行うため、負担をキャッシュ間に効率的に分散するのが困難である。つまり、負担を均等に分配しようとしているキャッシュが、負担が他のキャッシュ群に独自にどのように分配されているかを知り得ない場合があり、この結果、データオブジェクトのリクエストを1つ以上のキャッシュに予想外に集中させてしまう。 $\alpha = 1$ を選択しても、選択されたキャッシュまたは該キャッシュへの接続が一時的に作動不能になっていて、リクエストされたオブジェクトを該キャッシュが送信できない場合があるために効率的でない。従って、有利には「悲観的アプローチ」と呼ばれる方式が実行され、この方式では、決定関数の実行中はディレクトリリストは更新されず、ただしキャッシュが新しいオブジェクトを受けとるとディレクトリリストに適切なアドレスが加えられ、キャッシュがオブジェクトを排出（削除）するとそのアドレスがディレクトリリストから削除される。

【0046】 $\alpha = 1$ で、かつオブジェクトリクエストに対する応答が否定的な場合（UDP pingミス応答が受信されるか、または応答がない場合等）は、別のキャッシュが選択され、まったく新しいUDP pingメッセージの組がやりとりされて、リクエストされたオブジェクトを可能ならば他のキャッシュに送信させる。 $\alpha = 1$ を選択すると、ディレクトリリストが不正確で、選択された1つのキャッシュが実際はリクエストされたオブジェクトのコピーを持っていないという危険もある。 $\alpha > 1$ に選択することで、オブジェクトのコピー入手過程に冗長性をもたせ、リクエストされたオブジェクトを1つ以上の可能性のある場所で探索することによって、信頼性のないネットワーク中での上記の問題を軽減することができる。またこの冗長性により、悲観的アプローチ

チでは訂正されないいくつかの不正確なディレクトリエントリをもつ影響を緩和できる。

【0047】本発明の一実施形態では、 α に定数Kを選択する。この場合、ディレクトリリスト上のキャッシュ数NがK以下ならば、UDP pingリクエストはリクエストされたオブジェクトのコピーを記憶しているN個のキャッシュ、および該コピーを記憶していないK-N個のキャッシュへ送信され、本発明の効率は従来のキャッシングシステムのものに近づく。ディレクトリリスト上のキャッシュ数NがK以上ならば、キャッシュはディレクトリリストからK個のキャッシュを選び出し、各キャッシュにUDP pingリクエストを送る。この場合、作成されるメッセージ数は、リクエストされたオブジェクトのコピーを運ぶメッセージを除いて $2K + 3$ 以下となる。これは上記で $\alpha = 1$ について議論したのと同じ理由で信頼性のあるネットワーク中で有効となりうる。しかし信頼性のないネットワークでは、リクエストされたオブジェクトのコピーをもっていると称するディレクトリリスト上の全キャッシュ群にUDP pingリクエストを送信することによって得られる冗長性を完全には利用できない。

【0048】本発明に従えば、 α 値は、有利には本発明が実現されるネットワークの信頼性に従ってキャッシングシステムの性能を最大にするべく選択される。信頼性の高いネットワークでは、 α は1に近く選択される。信頼性の低いネットワークでは、 α は1より大きく選択され、ネットワークの信頼性が低いほどより大きな値が選択される。ただし信頼性のないネットワークでの α の値は、多くてもディレクトリリストのアドレス数と同じにしなければならない。信頼性のないネットワーク上で実現される本発明の一実施形態では、各ディレクトリリストのサイズはK個のアドレスに固定され、 α はKに設定される。本発明の他の実施形態では、 α は全ディレクトリリスト群の平均アドレス数に設定される。さらに他の実施形態では、 α は全ディレクトリリスト群のメジアン数に設定される。楽観的アプローチでは、あるキャッシュがオリジナルコピーを受けとった場合、すなわちオブジェクトのオリジナル情報源にオブジェクトのリクエストを送信した場合にのみ、キャッシュはオブジェクトのディレクトリリストを記憶しているキャッシュにディレクトリ追加リクエストを送信する。ディレクトリ追加リクエストを受信すると、ディレクトリキャッシュは、所与のオブジェクトのディレクトリリストに送信元のキャッシュのアドレスを追加する。また楽観的アプローチでは、所与のオブジェクトのコピーをリクエストするキャッシュがオブジェクトキャッシュからリクエストされたオブジェクトのコピーが受信できなかった場合は、リクエスト元のキャッシュはディレクトリキャッシュにディレクトリ削除リクエストを送信し、該オブジェクトのディレクトリリストからそのオブジェクトキャッシュのア

ドレスを削除する。また、オブジェクトキャッシュがオブジェクトを排出するたびに、オブジェクトキャッシュから適当なディレクトリキャッシュへディレクトリ削除リクエストが送信され、これによりオブジェクトキャッシュのアドレスがディレクトリストから削除される。

【0049】悲観的アプローチでは、キャッシングがどの情報源からでも所与のオブジェクトの新しいコピーを受信するたびに、キャッシングはディレクトリキャッシュにディレクトリ追加リクエストを送信する。リクエスト元のキャッシングがリクエストされたコピーが入手できなくとも、ディレクトリ削除リクエストは送信されない。ディレクトリ削除リクエストは、オブジェクトを記憶しているキャッシングがそのオブジェクトを排出するときにのみ適当なディレクトリキャッシュに送信される。このディレクトリ削除リクエストはオブジェクト排出元のキャッシングから送信される。削除リクエストを受信すると、オブジェクト排出元キャッシングのアドレスがディレクトリストから削除される。

【0050】図12に示すように、データオブジェクトは、古くなると、記憶されているキャッシングから排出されなければならない。これは本発明の一形態に従って、生存時間(TTL)パラメータを記憶されている各オブジェクトに関連づけることによって行われる。TTLパラメータはキャッシングされたオブジェクトの最大寿命を特定する。一実施形態では、TTLパラメータは、オブジェクトのコピーが受信されキャッシングに記憶された時からある一定の時間量だけれた日付スタンプである。TTLが期限切れすると、オブジェクトは排出される。TTLパラメータは従来のシステム中で実現されており、古くなったコピーをキャッシング中に残留させてしまうという欠点がある。従来のシステムでは、この問題は所与のオブジェクトが古くなる(オリジナルが変更される等)と、該オブジェクトのすべてのキャッシングに排出メッセージを配信することによって対処される。しかし、排出メッセージの配信は、ネットワークに不利に負担をかける大量のトラフィックを生じてしまう。

【0051】本発明は、古くなったキャッシングされたコピーを効率的に排出して、費用のかかる配信メッセージの使用を避けることによって、この問題を解決する。本発明に従えば、オリジナルオブジェクトが変更されるたびにオブジェクトのオリジナル情報源からディレクトリキャッシングに排出リクエストを送信することによって、オブジェクトをそのTTLが期限切れする前に排出できる。その後、ディレクトリキャッシングは、オブジェクトのディレクトリスト上の全オブジェクトキャッシングに排出メッセージを送信する。これを受けて、オブジェクトキャッシングはオブジェクトを排出し、ディレクトリスト 자체がディレクトリキャッシングから削除される。

【0052】この様子を図12に示す。キャッシングA1201は、オリジナル情報源サーバ1202からオブ

エクトYの全コピーの排出リクエストを受信する。キャッシングA1201はディレクトリロケータ関数L(Y)=C1203を実行し、オブジェクトYのディレクトリストがディレクトリキャッシングC1204に記憶されていることを知る。キャッシングA1201はキャッシングC1204にディレクトリ排出リクエストを送信する(direct_eject(1))。キャッシングC1204上のディレクトリストYは、YのキャッシングされたコピーがキャッシングE1205およびF1206に記憶されていることを示す。ディレクトリ排出リクエストを受信すると(direct_eject(2))、キャッシングC1204はキャッシングE1205およびF1206に排出リクエスト(eject(3))を送り、これらのキャッシングはYのコピーを排出する。その後、キャッシングC1204はそのYのディレクトリストを削除する。他の実施形態では、キャッシングAは、オブジェクトのオリジナル情報源サーバからオブジェクトのコピーを受信するたびに、ディレクトリサーバにディレクトリ排出リクエストを送信する。この実施形態はオリジナル情報源サーバからの排出メッセージを必要としない。本発明は、従来のいくつかのキャッシングシステムで実行された全ディレクトリキャッシングへの排出メッセージの配信を不要にできる。これにより古くなったオブジェクトのコピーの排出のために生じるネットワークトラフィック量を大幅に削減でき、キャッシングシステムの正確さをより効率的に維持できる。

【0053】本発明に従って実現される分散型排出方式は、当該技術分野で公知である全キャッシングへの排出メッセージの配信を不要にする。分散型排出方式は、古くなったオブジェクトの排出に必要なネットワークトラフィックだけしか生じないため、配信方式よりも効率的かつ経済的である。さらに、分散型排出は大規模なシステムにも小規模なシステムと同様に有効に動作するので、ネットワークのサイズが大きくなるとそれに応じて負担となる大量のトラフィックを生じる配信方式よりもスケーラブルである。

【0054】本発明は、従来のキャッシングシステムと比べて効率が高く、かつスケーラブルなキャッシングシステムを提供する。このシステムは、ディレクトリストの形の分散型ディレクトリを提供し、各ディレクトリストは1つのデータオブジェクトに関連づけられる。ディレクトリストは、オブジェクトのコピーが記憶されているオブジェクトキャッシング群のリストである。オブジェクトのディレクトリストはネットワークの多数のサーバ間に分散される。所与のオブジェクトのディレクトリストの位置は、オブジェクトのネットワークアドレスをディレクトリストが記憶されているディレクトリキャッシングのネットワークアドレスを指すポインタに相關させるハッシュロケータ関数によって、簡単かつ経済的に求められる。ディレクトリストの多数のオブ

ジェクトキャッシュ群がポーリングされ、その中からキャッシュされたコピーを取り出す1つのキャッシングが選択される。オブジェクトキャッシングの数は、信頼性の異なる様々なネットワーク環境でキャッシングシステムの性能を最大にするように、本発明に従って有利に選択できる。ディレクトリリストは、ネットワークの状態に従って分散式で更新される。古くなったオブジェクトおよびディレクトリは、効率的な分散方式でキャッシングシステムから排出される。本発明は、従来のキャッシングシステムに比べて高い信頼性と低い費用で実現することができる。さらに、本発明はボトルネックや单一故障点の影響を受けにくいので、従来のキャッシングシステムに関連した多くの問題を解決でき、ネットワークにかけるキャッシングシステムのネットワークトラフィックの負担を従来のシステムより大幅に削減できる。

【0055】以下に、本発明の様々な形態の仮想的なコードの実施例を示す。所与のオブジェクトのコピーの入手を望むクライアントは、あるキャッシングに入手リクエストメッセージを送信する。図5に示すように、オブジェクトYに対する入手リクエストはクライアント502からキャッシングA501に送られる。入手リクエストの仮想的なコードの実施例は図13のようになる。

【0056】上記の入手リクエストの実施例はディレクトリ探索リクエストを実行する。探索リクエストの仮想的なコードは図14のようになる。

【0057】ディレクトリリストから α 個のアドレスを選択するディレクトリ探索ルーチンの仮想的なコードの実施例は図15のようになる。

【0058】UDP pingリクエスト／ヒット／ミス・ルーチンの仮想的なコードの実施例は図16のようになる。

【0059】ディレクトリ追加ルーチンの仮想的なコードの実施例は図17のようになる。

【0060】ディレクトリ削除ルーチンの仮想的なコードの実施例は図18のようになる。

【0061】排出ルーチンの仮想的なコードの実施例は図19のようになる。

【0062】オリジナル情報源からオブジェクトのコピー入手するリフレッシュルーチンの仮想的なコードの実施例は図20のようになる。

【0063】応答が目的地から受信されたかどうかをテストし、適当な維持方式を実行するルーチンの仮想的なコードの実施例は図21のようになる。

【0064】本発明に従う分散型排出方式の仮想的なコードの実施例は図22のようになる。

【0065】排出ルーチンの仮想的なコードの実施例は図23のようになる。

【0066】本発明は、有利には、従来のキャッシングシステムより効率的な、オブジェクトのキャッシングシステムおよび方法をネットワーク上で提供する。本発明

に従う分散型方法は、有利には、キャッシングされたオブジェクトを記憶、発見、入手、および排出を行うスケーラブルな方法を提供し、ネットワーク上のデータオブジェクト数が増大しても、ネットワークに課すキャッシングシステムトラフィックの負担を低く抑えつつ性能を向上させることができる、改良されたネットワークを提供する。本発明はまた、データオブジェクトが頻繁に変化するネットワークにおいて、従来のキャッシングシステムよりも優れた性能を發揮する。これは、有利には従来のキャッシングシステムにおいて不都合にネットワークに負担をかけていたロッキング、承認、および配信メッセージを必要とすることなく達成できる。

【図面の簡単な説明】

【図1】先行技術のキャッシングのシステムの実施形態を示す図である。

【図2】先行技術のキャッシングのシステムにおける UDP pingリクエストの伝達過程を示す図である。

【図3】先行技術のキャッシングのシステムにおける UDP pingヒットメッセージおよび UDP pingミスマッチメッセージの伝達過程を示す図である。

【図4】先行技術のキャッシングのシステムにおいて、オブジェクトのコピーに対するリクエストを伝達し、リクエストされたオブジェクトのコピーをキャッシングおよびクライアントに伝達する過程を示す図である。

【図5】本発明に従うキャッシングのシステムの一実施形態において、クライアントがオブジェクトに対するリクエストを受信キャッシングに送信し、受信キャッシングがリクエストされたオブジェクトに対してディレクトリロケータ関数を実行する様子を示す図である。

【図6】本発明に従うキャッシングのシステムの一実施形態において、クライアントからのオブジェクトに対するリクエストに応答して、受信キャッシングがディレクトリキャッシングに探索リクエストを送信する様子を示す図である。

【図7】本発明に従うキャッシングのシステムの一実施形態において、ディレクトリキャッシングが、リクエストされたオブジェクトのディレクトリリスト上のキャッシング群に UDP pingリクエストを送信する様子を示す図である。

【図8】本発明に従うキャッシングのシステムの一実施形態において、リクエストされたオブジェクトのディレクトリリスト上のキャッシング群が、ディレクトリキャッシングに対して UDP pingヒットメッセージで応答する様子を示す図である。

【図9】本発明に従うキャッシングのシステムの一実施形態において、受信キャッシングがオブジェクトリクエストメッセージをディレクトリリスト上のキャッシングに送信し、このキャッシングがリクエストされたオブジェクトのコピーを受信キャッシングに送信し、受信キャッシング

がオブジェクトのコピーをクライアントに送信する様子を示す図である。

【図10】 本発明に従うキャッシングのシステムの一実施形態において、オブジェクトキャッシュ群の1つのキャッシュにだけ、オブジェクトリクエストメッセージが送信される様子を示す図である。

【図11】 本発明に従うキャッシングのシステムの一実施形態において、キャッシングがディレクトリキャッシングに削除リクエストメッセージを送信する様子を示す図である。

【図12】 本発明に従うキャッシングのシステムの一実施形態において、サーバが受信キャッシングに所与のオブジェクトに対するディレクトリ排出メッセージを送信し、受信キャッシングがこのオブジェクトに対してディレクトリロケータ関数を実行し、かつディレクトリキャッシングにディレクトリ排出メッセージを送信し、ディレクトリキャッシングがオブジェクトに対してディレクトリスト上のキャッシングに排出メッセージを送信する様子を示す図である。

【図13】 入手リクエストの仮想コードを示す図である。

【図14】 探索リクエストの仮想コードを示す図である。

【図15】 ディレクトリ探索ルーチンの仮想コードを示す図である。

【図16】 pingリクエストの仮想コードを示す図である。

【図17】 ディレクトリ追加ルーチンの仮想コードを示す図である。

【図18】 ディレクトリ削除ルーチンの仮想コードを示す図である。

【図19】 排出ルーチンの仮想コードを示す図である。

【図20】 リフレッシュルーチンの仮想コードを示す図である。

【図21】 レスポンス待ちルーチンの仮想コードを示す図である。

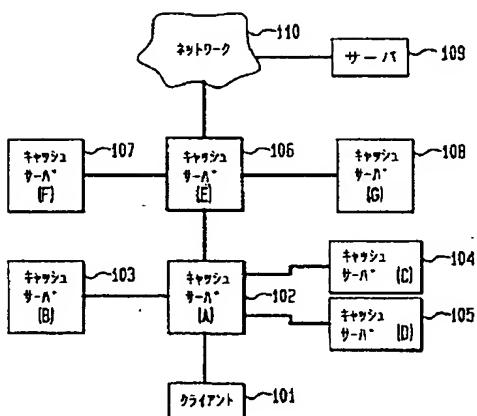
【図22】 ディレクトリ排出ルーチンの仮想コードを示す図である。

【図23】 排出ルーチンの仮想コードを示す図である。

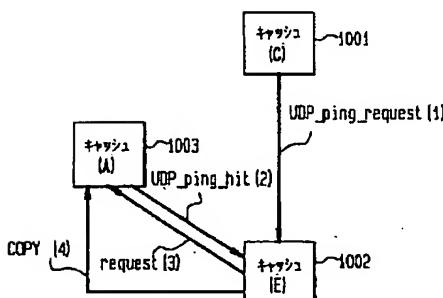
【符号の説明】

- 501 受信キャッシング、502 クライアント、503 ディレクトリキャッシング、504, 505, 506, 507 オブジェクトキャッシング、508ハッシュロケータ関数。

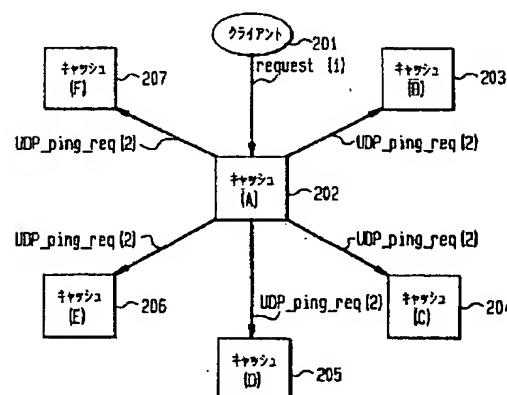
【図1】



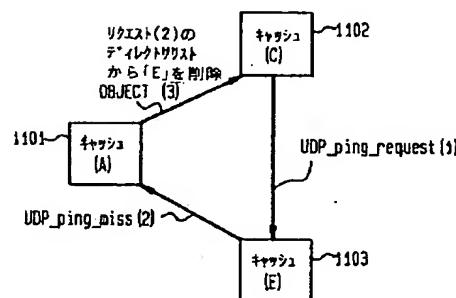
【図10】



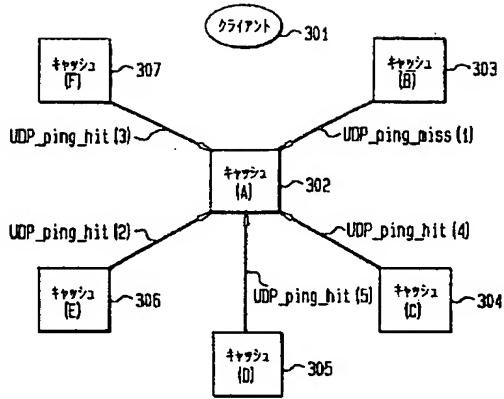
【図2】



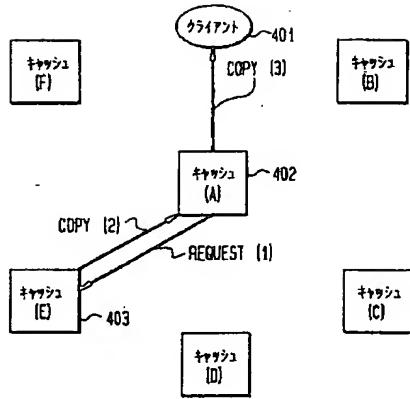
【図11】



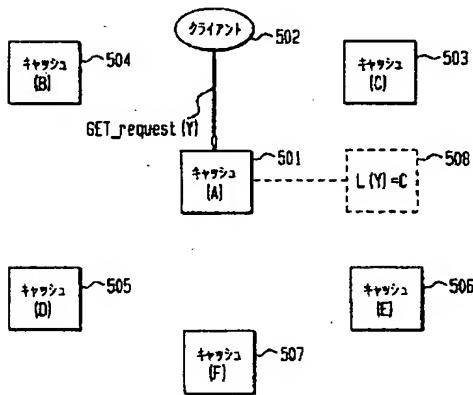
【図3】



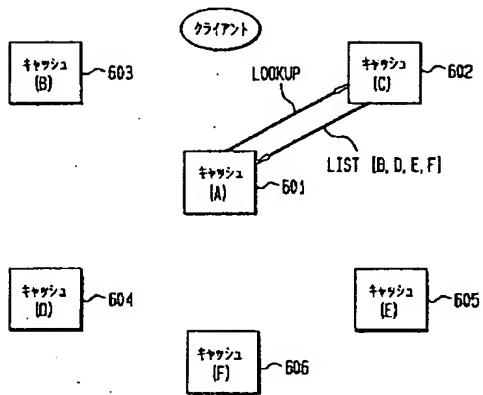
【図4】



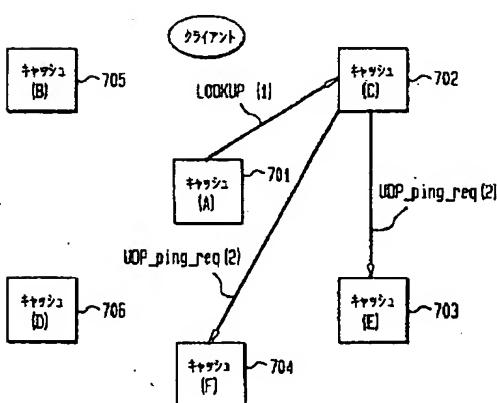
【図5】



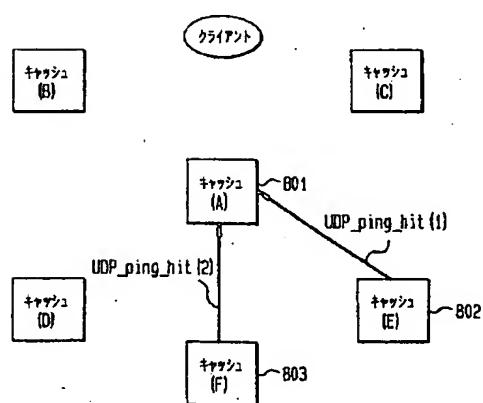
【図6】



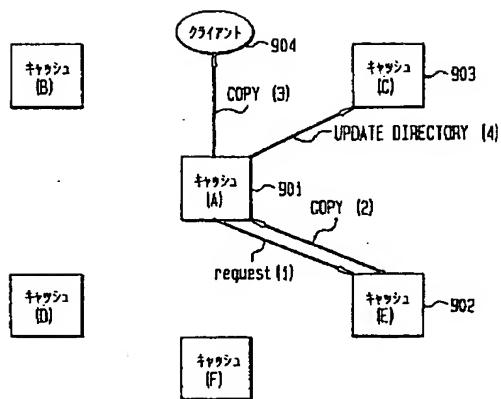
【図7】



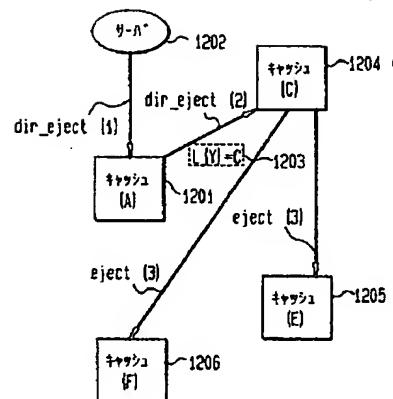
【図8】



【図9】



【図12】



【図14】

```

dir_lookup (input:url_of_object, client_IPaddr, output:list_of_neighbors )
/*client は探索リクエストの送信元を指す */
{
    find_URL(url_of_object);
    /* リクエストされたオブジェクトのネットワークアドレスに関連
     したディレクトリリストを見つける */

    if URL_of_object is found
    then return(list_of_neighbors);
    /*もしリクエストされたオブジェクトのディレクトリリストが見
     つかれば、ディレクトリリスト上のアドレスリストを返送 */

    else return(null_list);

    if optimistic approach
    find_CLIENT(client_IPaddr);
    /* 乐观的アプローチにおいて、もし顧客キャッシュアドレスがま
     たディレクトリリスト上になければ、顧客キャッシュのアドレ
     スを加えてディレクトリリストが更新される */

    if client_IPaddr is not found
    add_neighbor_to_list(url_of_object, client_IPaddr);
}
  
```

【図23】

```

eject (input:url_of_object)
{
    find object url_of_object in local cache;
    if object is found
    then remove object url_of_object from the local cache;
}
  
```

[図13]

```

GET_request (input: url_of_object )
/* url_of_object はリクエストされたオブジェクトのネットワーク
アドレス */
{
...
dest_IPAddr:=HASH_TABLE_NEIGHBORS[hash_1(url_of_object)];
/* この行は本実施例でオブジェクトのネットワークアドレスに
対して実行されるハッシュ関数であるディレクトリロケータ
関数を示す */
if dest_IPAddr is equal to my_IPAddr
    /* ロケータ関数がディレクトリキャッシュ自身を指し
    ているか？ */
then lookup(url_of_object, my_IPAddr);
    /* ロケータ関数がディレクトリキャッシュを指して
    いれば、リクエスト元の顧客にコピーを送信する
    探索ルーチンが実行される */
else send(dest_IPAddr, lookup(url_of_object, my_IPAddr));
    /* この行は探索関数で示されたディレクトリリストが位置する、
    受信キャッシュ以外のキャッシュ（「ディレクトリキャッシュ」）
    に探索リクエストを送信する */
loop until(response is equal to UDP_ping_hit
            or response is equal to NO_neighbors
            or waiting_time expired
            or number_of_responses is equal to alpha)
    wait(response, neighbor_IPAddr);
    /* このループは、最初の UDP ping 当たりメッセージが受信される
    まで、または受信キャッシュが直接通信可能なキャッシュに
    リクエストされたオブジェクトのコピーが記憶されていないと
    いうメッセージになるまで、または所与の待機時間が応答なし
    で時間切れするまで、または応答数が所定数  $\alpha$  になるまで続行
    する */
if response is UDP_ping_hit
then send(neighbor_IPAddr, request(url_of_object, my_IPAddr));
/* UDP ping ヒットを受信すると、リクエストされたオブジェクトを
ネットワークアドレス my_IPAddr の受信キャッシュに送信するよう
に UDP ping ヒットの送信側にメッセージが送信される */
else send(home_url_IPAddr, Get_request(url_of_object, my_IPAddr));
    /* もし UDP ping リクエストが受信されなければ、home_url_IPAddr の
    リクエストされたオブジェクトの元リースが、リクエストされ
    たオブジェクトのコピーを受信キャッシュに送信するよう依頼
    される */
...
}

```

【図15】

```

lookup (input:url_of_object, requester_IPAddr )
{
dir_lookup(url_of_object, requester_IPAddr, list_of_neighbors);

if list_of_neighbors is null_list
then send(requester_IPAddr, (NO_neighbors, my_IPAddr));
else
make alpha_list by choosing alpha neighbors from list_of_neighbors;
send(alpha_list, UDP_ping_req (requester_IPAddr));
}

```

【図16】

```

UDP_ping_req (input:url_of_object, client_IPAddr)
/*client は UDP ping request の送信元を指す */
{
find object url_of_object in local cache;

if object is found
then send(client_IPAddr, UDP_ping_hit (my_IPAddr));
else send(client_IPAddr, UDP_ping_miss (my_IPAddr));
}

```

【図17】

```

dir_add (input:url_of_object, client_IPAddr )
/*client は オブジェクトのコピーが加えられたキャッシュを指す */
{
find_URL (url_of_object);
if url_of_object is not found
add_new_list (url_of_object); /*新しいディレクトリを開始*/
find_CLIENT (client_IPAddr);
/*ディレクトリリスト上の顧客ネットワークアドレスを探す*/
if client_IPAddr is not found
add_neighbor_to_list (url_of_object, client_IPAddr);
/*もし見つからなければこれをディレクトリリストに追加*/
}

```

【図18】

```

dir_del (input:url_of_object, client_IPAddr )
/*client はオブジェクトのコピーが削除されているキャッシュを
指す*/
{
find_URL (url_of_object);
/*ディレクトリリスト上のオブジェクトのアドレスを探す*/
if url_of_object is found
/*アドレスが見つかれば、そのアドレスをディレクトリリストか
ら削除*/
find_CLIENT (client_IPAddr);
if client_IPAddr is found
delete_neighbor_from_list (url_of_object, client_IPAddr);
}

```

【図19】

```

eject_local_copy (input:url_of_object )
{
...
dest_IPAddr:=HASH_TABLE_NEIGHBORS [hash_1(url_of_object)];
/*この行はオブジェクトのディレクトリリストが記憶されている
キャッシュの場所を突き止めるためのハッシュ関数としてディ
レクトリロケータ関数を実行する*/

if dest_IPAddr is equal to my_IPAddr
/*ディレクトリキャッシュがローカルキャッシュならば、ローカ
ルキャッシュのディレクトリリストを削除*/
then dir_del(url_of_object, my_IPAddr);
else send(dest_IPAddr, dir_del(url_of_object, my_IPAddr));
/*これ以外の場合はディレクトリ削除リクエストをディレクトリ
キャッシュに送信*/
}

```

【図20】

```

refresh (input:url_of_object )
{
...
dest_IPAddr:=HASH_TABLE_NEIGHBORS [hash_1(url_of_object)];
/*ディレクトリロケータ関数*/
if optimistic approach
/*楽観的アプローチでは、リフレッシュルーテンが実行されると
ディレクトリリストを更新する*/
if dest_IPAddr is equal to my_IPAddr
then dir_add(url_of_object, my_IPAddr);
else send(dest_IPAddr, dir_add(url_of_object, my_IPAddr));
}

```

【図21】

```

wait_response(input:status)
{
  ..
dest_IPAddr:=HASH_TABLE_NEIGHBORS[hash_1(url_of_object)];
/* この行はディレクトリロケータ関数である */

if status is OK and pessimistic approach
/* 悲観的アプローチにおいてコピーが受信されていれば、受信キ
   ャッシュのアドレスを加えてディレクトリストが更新される */

if dest_IPAddr is equal to my_IPAddr
then dir_add(url_of_object, my_IPAddr);
else send(dest_IPAddr, dir_add(url_of_object, my_IPAddr));

if status is not OK and optimistic approach
/* 楽観的アプローチにおいてコピーが受信されていなければ、
   受信キッシュのアドレスを削除してディレクトリストが
   更新される */
if dest_IPAddr is equal to my_IPAddr
then dir_del(url_of_object, my_IPAddr);
else send(dest_IPAddr, dir_del(url_of_object, my_IPAddr));
}

```

【図22】

```

dir_eject(input:url_of_object )
{
dest_IPAddr:=HASH_TABLE_NEIGHBORS[hash_1(url_of_object)];
/* この行はディレクトリ探索関数である */

if dest_IPAddr is not equal to my_IPAddr
/* まずローカルキッシュをチェックする */
then send(dest_IPAddr, dir_eject(url_of_object));
else
  find_URL(url_of_object); /* ディレクトリストを見つける */
  if url_of_object is found
    then
      send(list_of_neighbors, eject(url_of_object));
      clear the list_of_neighbors; /* ディレクトリストをクリアする */

    eject(url_of_object); /* ローカルコピーを排出する */
}

```

フロントページの続き

(72)発明者 パーサー ピー ドゥッタ
 アメリカ合衆国 カリフォルニア州 サン
 ノゼ マリブ ドライブ 1164

(72)発明者 トマス ピー ロンドン
 アメリカ合衆国 カリフォルニア州 マウ
 ンテン ビュー ラモス コート 2739

(72)発明者 シニサ サーブルジック
クロアチア ヴァリカ ゴリカ 10410
ジョシパ ジェラシオス 91 ヴェーネ
ュー 1

(72)発明者 ダリボー エフ ヴルサロヴィック
アメリカ合衆国 カリフォルニア州 サニ
ーヴェイル カムサック コート 932